

semhas



Model *Hierarchical Bayes* pada *Small Area Estimation* untuk Pendugaan Proporsi Pengangguran pada Desain Survey Kompleks

Arip Juliyanto
NRP.1314201712

PROGRAM MAGISTER
JURUSAN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU
PENGETAHUAN ALAM
INSTITUT TEKNOLOGI
SEPULUH NOPEMBER
SURABAYA
2015

Pembimbing:
Dr.rer.pol. Heri Kuswanto,M.Si
Dr.Dra. Ismaini Zain,M.Si

Pokok Bahasan



Pendahuluan

- Latar Belakang
- Rumusan Masalah
- Tujuan
- Manfaat
- Batasan



Tinjauan Pustaka

- Kebutuhan SAE
- SAE
- Konsep Bayes pada SAE
- Model Bayes SAE
- HB Area Level proporsi
- Evaluasi dan Perbandingan Model
- Two Stage Sampling
- Pengangguran



Metodologi

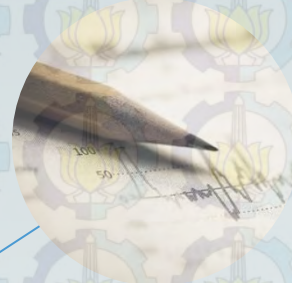
- Sumber data
- Variabel Penelitian
- Definisi operasional
- Metode Analisis
- Tahap Penelitian
- Jadwal kegiatan

Pendahuluan

BATASAN PENELITIAN



LATAR BELAKANG



TUJUAN PENELITIAN



RUMUSAN MASALAH





Pendahuluan (1)

Latar Belakang

- ✓ Pengangguran-Indikator penting
- ✓ Rata-rata pengangguran Riau sebesar 8,98 (2004-2013) terbesar kedua se Pulau Sumatera
- ✓ Informasi pengangguran masih terbatas sampai tingkat kabupaten

Pengangguran
Kecamatan

Sampel tidak
memadai

Solusi

Menambah
sampel survey

Sumberdaya
yang besar

Optimalisasi data
yang tersedia

SAE (Small Area
Estimation)

Pendahuluan (2)

Latar Belakang (SAE)



Small Area Estimation
(SAE)

BLUP

EB

HB

Dari beberapa penelitian [Ghosh dan Rao (1994), Arora dan Lahiri (1997), dan You dan Chapman (2006)] menyimpulkan bahwa SAE dengan pendekatan HB menghasilkan pendugaan yang relatif lebih baik

Pendahuluan (3)

Latar belakang (Penelitian terdahulu)

Xia, Carlin, dan Waller (1995) yang menggunakan *HB* untuk pemetaan rata-rata penderita kanker paru-paru di daerah Ohio pada tahun 1988.

Xie, Raghunathan, dan Lepkowski (2006) menggunakan metode *HB* untuk mengestimasi proporsi penduduk dengan berat badan yang tinggi di Amerika Serikat pada tahun 2003.

Pertiwi (2012) menggunakan *HB* untuk kepentingan menduga pengeluaran per kapita wilayah kecil.

Ubaidillah (2014) yang menggunakan model dua level dengan modifikasi pemodelan pada level area dengan pendekatan *Neural Network (NN)*.

You dan Rao (2002) dengan *HB unmatched sampling* dan *linking model* untuk pendugaan jumlah penduduk yang tidak tercakup dalam sensus.


Mohadjer, Rao, Liu, Krenzke dan Kerckhove (2005) menggunakan *unmatched sampling* dan *linking model* untuk estimasi tingkat kemampuan baca-tulis di Amerika.




Pendahuluan (4)

Latar Belakang (Asumsi SRS dan design consistent)

- ✓ Pemodelan SAE seringkali mengasumsikan data berasal dari desain survey SRS (simple random sampling)
- ✓ Data publik seringkali berasal dari desain survey yang lebih kompleks dari SRS (stratifikasi, two stage sampling. Two phase dll)



Liu *et al.* (2007) mempertimbangkan desain survey stratifikasi SRS pada empat model HB baik dengan asumsi varian diketahui maupun tidak.



Ha *et al.* (2014) dengan menggunakan data dari NHIS menduga proporsi perokok dengan pertimbangan desain survey yang lebih kompleks (*two stage sampling survey*) dengan menggunakan tiga jenis model HB

Design
consistent

Pendahuluan (4)

- Tipe Small Area estimation (SAE)



Unit level model

Area level model



- Variabel covariate jarang tersedia
- Design consistent sulit terpenuhi

- Variabel covariate lebih banyak tersedia
- Design consistent

Pendahuluan (5)

Rumusan Masalah

- ✓ Bagaimana penerapan beberapa model HB dalam menduga proporsi pengangguran tingkat kecamatan pada desain survey kompleks (*two stage sampling*).
- ✓ Bagaimana perbandingan beberapa model HB (*FH, NL, NLrs, BLrs*) dalam pendugaan proporsi pengangguran.

Tujuan

- Menerapkan beberapa model HB untuk menduga proporsi pengangguran di tingkat kecamatan pada desain survey kompleks (*two stage sampling*).
- Membandingkan performa dari beberapa model HB dalam pendugaan proporsi pengangguran pada desain survey kompleks (*two stage sampling*).

Pendahuluan (6)

Manfaat

- ✓ Memberikan informasi mengenai performa beberapa model HB SAE dengan berbagai model estimasi .
- ✓ Sebagai bahan diskusi selanjutnya untuk pemodelan yang lebih tepat .
- ✓ Tersedianya data atau indikator proporsi pengangguran pada tingkat kecamatan.

Batasan

- Konsep BPS (Sakernas dan SP2010).
- Desain sampling yang digunakan *two stage sampling* PPS-SRS.
- Model HB yang digunakan adalah yang diajukan oleh Liu *et al.* (2007) dan Ha *et al.* (2014) dengan berbagai asumsi pada level sampling.
- Model SAE yang digunakan menggunakan model *area level*.

TINJAUAN PUSTAKA (1)

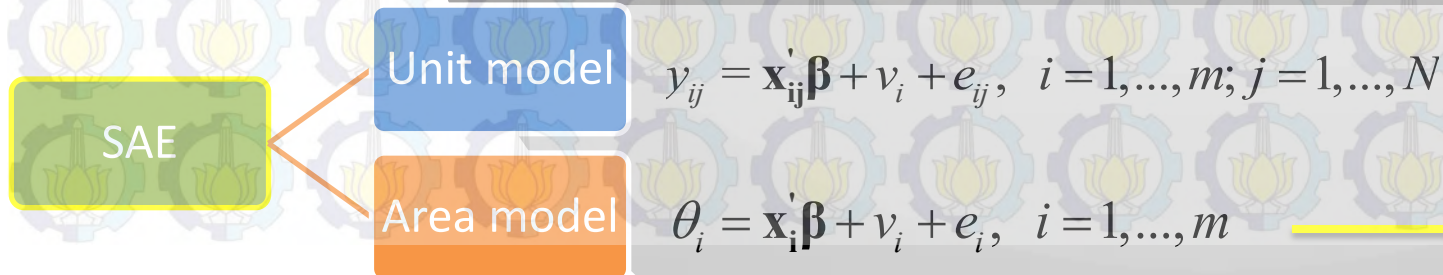


Kebutuhan SAE

- Meningkatnya tuntutan akan small area statistik
- Derasnya arus informasi
- Ketersediaan dana (Setiawan, 2004)

Small Area Estimation (SAE)

- Metode SAE estimasi tidak langsung (*indirect estimation*) yang bersifat ‘meminjam’ kekuatan (*borrowing strenght*) dari pengamatan sampel area yang berdekatan dengan memanfaatkan informasi tambahan (Rao, 2003)



TINJAUAN PUSTAKA (2)



Model SAE HB Fay-Herriot (F-H)

- Menjadi model area SAE yang seringkali digunakan untuk pendugaan parameter proporsi (Liu, 2009)

Level 1 (*sampling model*) : $\hat{\theta}_{iw} | \theta_i \overset{ind}{\sim} N(\theta_i, \psi_i)$

Level 2 (*linking model*) : $\theta_i | \boldsymbol{\beta}, \sigma_v^2 \overset{ind}{\sim} N(\mathbf{x}'_i \boldsymbol{\beta}, \sigma_v^2)$

ψ_i is sampling variance (diasumsikan diketahui)

\mathbf{x} vector covariate variabel

$\hat{\theta}_{iw}$ proporsi observasi

θ_i paramater proporsi yang hendak diduga

Dapat dibentuk menjadi persamaan Linier Mixed Model (Model Linier Campuran)

$$\hat{\theta}_{iw} = \mathbf{x}'_i \boldsymbol{\beta} + v_i + \varepsilon_i$$

TINJAUAN PUSTAKA (3)



Model Normal-Logistic (NL)

- Asumsi normalitas pada linking model menjadi kurang baik terutama ketika 'true' proporsi mendekati 0 atau 1
- Asumsi ini juga tidak menjamin posterior P selalu jatuh pada range 0 dan 1

Level 1 (*sampling model*) $:\hat{\theta}_{iw}|\theta_i \stackrel{ind}{\sim} N(\theta_i, \psi_i)$

Level 2 (*linking model*) $: \text{logit}(\theta_i)|\boldsymbol{\beta}, \sigma_v^2 \stackrel{ind}{\sim} N(\mathbf{x}'_i\boldsymbol{\beta}, \sigma_v^2)$

ψ_i is sampling variance (diasumsikan diketahui)

Model FH dan NL, nilai ψ_i diestimasi dengan (Ha, et al. 2013):

$$\hat{\psi}_i = \frac{\hat{\theta}_{j(i)}^{rgn} (1 - \hat{\theta}_{j(i)}^{rgn})}{n_i} deff_{iw}$$

$$deff_{iw} = \frac{\text{var}_{dsgn}(\hat{\theta}_{j(i)}^{rgn})}{\text{var}_{srs}(\hat{\theta}_{j(i)}^{rgn})}$$

TINJAUAN PUSTAKA (4)



Model Normal-Logistic Random Sampling (NLrs)

- Model Fay-Herriot dan NL di atas mengasumsikan varian sampling diketahui, meskipun demikian tetap harus diestimasi
- Sebagai alternatif, model NLrs mengasumsikan bahwa varian sampling tidak diketahui

Level 1 (sampling model) : $\hat{\theta}_{iw} | \theta_i \stackrel{ind}{\sim} N(\theta_i, \psi_i),$

Level 2 (linking model) : $\text{logit}(\theta_i) | \boldsymbol{\beta}, \sigma_v^2 \stackrel{ind}{\sim} N(\mathbf{x}'_i \boldsymbol{\beta}, \sigma_v^2)$

ψ_i is sampling variance (diasumsikan TIDAK diketahui)

$$\psi_i = \left(\frac{\theta_i(1-\theta_i)}{n_i} \right) * deff_{iw}$$

TINJAUAN PUSTAKA (5)



Model Beta-Logistic Random Sampling (BLRs)

- Untuk mengakomodasi distribusi yang tidak normal (asymetris), Hawala dan Lahiri (2012) serta Liu, et al. (2013) mengusulkan pendekatan distribusi Beta pada sampling model
- Distrib Beta juga menggaransi range antara 0 dan 1 sebagaimana sifat proporsi

Level 1 (sampling model) : $\hat{\theta}_{iw} | \theta_i \stackrel{ind}{\sim} \text{Beta}(a_i, b_i)$,

Level 2 (linking model) : $\text{logit}(\theta_i) | \boldsymbol{\beta}, \sigma_v^2 \stackrel{ind}{\sim} N(\mathbf{x}'_i \boldsymbol{\beta}, \sigma_v^2)$

$$a_i = \theta_i \left(\frac{n_i}{deff_{iw}} - 1 \right) \quad b_i = (1 - \theta_i) \left(\frac{n_i}{deff_{iw}} - 1 \right)$$

TINJAUAN PUSTAKA (6)



Full Conditional Distribution

$$\left[\theta_i \mid \beta, \sigma_v^2, \tilde{\theta} \right] \sim N \left((1-r_i) \hat{\theta}_{iw} + r_i \mathbf{x}_i' \beta, \psi_i (1-r_i) \right), \theta_i \in R$$

$$\left[\beta \mid \theta_i, \sigma_v^2, \tilde{\theta} \right] \sim N \left(\left(\sum_{i=1}^m \mathbf{x}_i \mathbf{x}_i' \right)^{-1} \left(\sum_{i=1}^m \mathbf{x}_i \theta_i \right), \sigma_v^2 \left(\sum_{i=1}^m \mathbf{x}_i \mathbf{x}_i' \right)^{-1} \right), \beta \in \mathbf{R}^p$$

$$\left[\sigma_v^2 \mid \beta, \theta_i, \tilde{\theta} \right] \sim \begin{cases} ING \left(\frac{1}{2} m - 1, \frac{1}{2} \sum_{i=1}^m (\theta_i - \mathbf{x}_i' \beta)^2 \right), & \sigma_v^2 \in (0, L) \\ 0, & \sigma_v^2 \geq L \end{cases}$$

FH

$$\left[\theta_i \mid \beta, \sigma_v^2, \tilde{\theta} \right] \propto \frac{1}{\theta_i (1-\theta_i) \sigma_v \sqrt{\psi_i}} \exp \left(-\frac{(\hat{\theta}_{iw} - \theta_i)^2}{2\psi_i} - \frac{(\text{logit}(\theta_i) - \mathbf{x}_i' \beta)^2}{2\sigma_v^2} \right), \theta_i \in (0,1)$$

$$\left[\beta \mid \theta_i, \sigma_v^2, \tilde{\theta} \right] \sim N \left(\left(\sum_{i=1}^m \mathbf{x}_i \mathbf{x}_i' \right)^{-1} \left(\sum_{i=1}^m \mathbf{x}_i \text{logit}(\theta_i) \right), \sigma_v^2 \left(\sum_{i=1}^m \mathbf{x}_i \mathbf{x}_i' \right)^{-1} \right), \beta \in \mathbf{R}^p$$

$$\left[\sigma_v^2 \mid \beta, \theta_i, \tilde{\theta} \right] \sim \begin{cases} ING \left(\frac{1}{2} m - 1, \frac{1}{2} \sum_{i=1}^m (\text{logit}(\theta_i) - \mathbf{x}_i' \beta)^2 \right), & \sigma_v^2 \in (0, L) \\ 0, & \sigma_v^2 \geq L \end{cases}$$

NL

$$\left[\theta_i \mid \beta, \sigma_v^2, \tilde{\theta} \right] \propto \frac{1}{\theta_i (1-\theta_i) \sigma_v \sqrt{\psi_i}} \exp \left(-\frac{(\hat{\theta}_{iw} - \theta_i)^2}{2\psi_i} - \frac{(\text{logit}(\theta_i) - \mathbf{x}_i' \beta)^2}{2\sigma_v^2} \right)$$

$$\left[\beta \mid \theta_i, \sigma_v^2, \tilde{\theta} \right] \sim N \left(\left(\sum_{i=1}^m \mathbf{x}_i \mathbf{x}_i' \right)^{-1} \left(\sum_{i=1}^m \mathbf{x}_i \text{logit}(\theta_i) \right), \sigma_v^2 \left(\sum_{i=1}^m \mathbf{x}_i \mathbf{x}_i' \right)^{-1} \right)$$

$$\left[\sigma_v^2 \mid \beta, \theta_i, \tilde{\theta} \right] \sim \begin{cases} ING \left(\frac{1}{2} m - 1, \frac{1}{2} \sum_{i=1}^m (\text{logit}(\theta_i) - \mathbf{x}_i' \beta)^2 \right), & \sigma_v^2 \in (0, L) \\ 0, & \sigma_v^2 \geq L \end{cases}$$

NL
rs

$$\left[\theta_i \mid \beta, \sigma_v^2, \tilde{\theta} \right] \propto \frac{1}{\theta_i (1-\theta_i) \sigma_v} \frac{\hat{\theta}_{iw}^{\theta_i \delta_{iw} - 1} (1 - \hat{\theta}_{iw})^{(1-\theta_i) \delta_{iw} - 1}}{\Gamma(\theta_i \delta_{iw}) \Gamma(1-\theta_i) \delta_{iw}} \exp \left(-\frac{(\text{logit}(\theta_i) - \mathbf{x}_i' \beta)^2}{2\sigma_v^2} \right), \theta_i \in (0,1)$$

$$\left[\beta \mid \theta_i, \sigma_v^2, \tilde{\theta} \right] \sim N \left(\left(\sum_{i=1}^m \mathbf{x}_i \mathbf{x}_i' \right)^{-1} \left(\sum_{i=1}^m \mathbf{x}_i \text{logit}(\theta_i) \right), \sigma_v^2 \left(\sum_{i=1}^m \mathbf{x}_i \mathbf{x}_i' \right)^{-1} \right), \beta \in \mathbf{R}^p$$

$$\left[\sigma_v^2 \mid \beta, \theta_i, \tilde{\theta} \right] \sim \begin{cases} ING \left(\frac{1}{2} m - 1, \frac{1}{2} \sum_{i=1}^m (\text{logit}(\theta_i) - \mathbf{x}_i' \beta)^2 \right), & \sigma_v^2 \in (0, L) \\ 0, & \sigma_v^2 \geq L \end{cases}$$

BL
rs

TINJAUAN PUSTAKA (7)



Perbandingan Model

- Salah satu ukuran kecocokan model yang dapat digunakan dalam mengevaluasi sekaligus membandingkan beberapa model bayes adalah *Deviance Information Criterion* (DIC) sebagaimana yang diusulkan oleh Spiegelhalter et al. (2002).

$$DIC = 2E_{\theta|y}(D) - D[E_{\theta|y}(\theta)]$$

$$DIC = \bar{D} + p\hat{D} = 2\hat{D} - D(\hat{\theta})$$

$$\bar{D} = -2 \sum_{k=1}^K \log(y|\theta^k) / G \quad p\hat{D} = \bar{D} - D(\hat{\theta})$$

$$D(\hat{\theta}) = -2 \log(y|\hat{\theta})$$

Liu (2009) membandingkan beberapa kebaikan model SAE HB dengan menggunakan bias statistik

$$OAB = \frac{1}{mR} \sum_{i=1}^m \sum_{r=1}^R [\hat{\theta}_i^r - \theta_i]$$

$$OAAD = \frac{1}{mR} \sum_{i=1}^m \sum_{r=1}^R |\hat{\theta}_i^r - \theta_i|$$

$$OAARD = \frac{1}{mR} \sum_{i=1}^m \sum_{r=1}^R \frac{|\hat{\theta}_i^r - \theta_i|}{\theta_i}$$

TINJAUAN PUSTAKA (8)



Two Stage Sampling

- Salah satu dari kompleks survey
- Dalam penelitian ini dengan konsep EPSEM (*Equal Probability of Selection Method*) dengan skenario
Tahap 1. Dipilih BS (blok sensus) secara PPS dengan size jumlah RT
Tahap 2. Dari BS terpilih, dipilih 10 RT secara SRS WOR

Primary sampling unit : Blok Sensus
Sampling unit: Rumah Tangga
Elementary Unit: Anggota RT

TINJAUAN PUSTAKA (9)



Pengangguran Terbuka

- Tingkat pengangguran: banyaknya jumlah angkatan kerja yang tidak bekerja dan aktif mencari pekerjaan



TINJAUAN PUSTAKA (10)



Faktor Mempengaruhi /berkorelasi terhadap Pengangguran

- ✓ Sektor informal mempunyai peranan yang signifikan sebagai penyangga (*buffer*) dalam perekonomian dan dianggap menjadi penyebab tidak begitu signifikannya peningkatan angka pengangguran pada masa krisis (Prayitno, 2002).
 - ✓ Tenaga kerja yang berlebih di pedesaan memungkinkan untuk migrasi dari kemiskinan dan pengangguran dengan adanya sektor informal (Widodo, 2005)
 - ✓ Sari (2011) menyebutkan bahwa pengangguran regional di Indonesia mempunyai mekanisme yang lambat menuju keseimbangan disebabkan karena beberapa faktor yang berpengaruh terhadap tingkat pengangguran regional antara lain angkatan kerja berusia muda (15-24 tahun), angkatan kerja berpendidikan tinggi, angkatan kerja yang berjenis kelamin pria, *dependency ratio*, dan komposisi industri
 - ✓ Karoma (2013) dalam *Indonesian Palm Oil Magazine* (2013) menyatakan bahwa pertanian adalah solusi bagi masalah pengangguran di Sierre Leone yang memiliki tingkat pengangguran kaum muda tertinggi di Afrika
- ❖ Sektor informal
 - ❖ Pendidikan
 - ❖ Rasio jenis kelamin
 - ❖ Komposisi industri
 - ❖ Bantuan Usaha

Metodologi (1)

Sumber data

Untuk simulasi model SAE HB

- Komponen penduga proporsi pengangguran berasal dari hasil Sensus 2010 (SP2010)
- Variable covariate berasal dari data Podes 2011

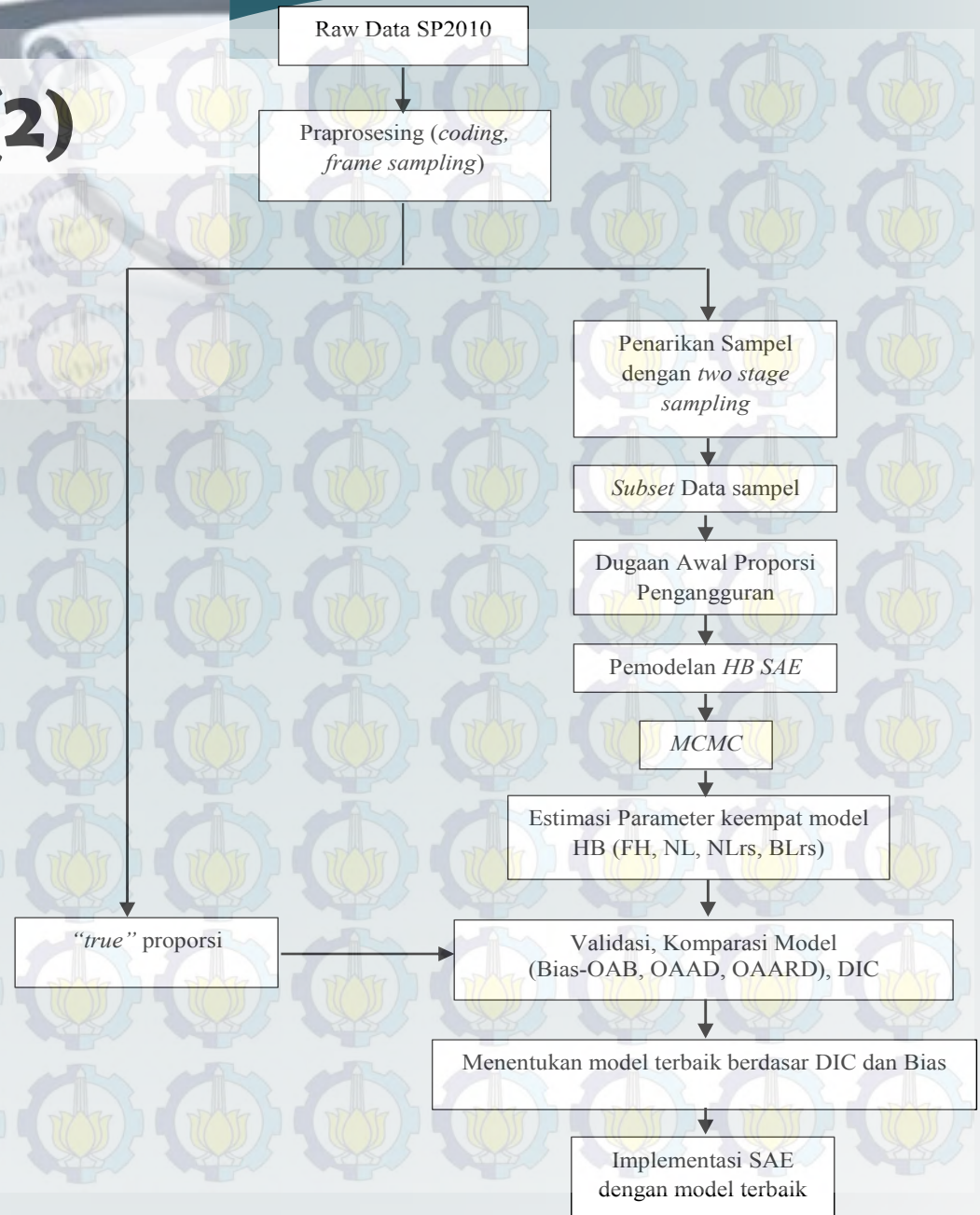
Untuk implementasi model menggunakan raw data sakernas (hasil entry) 2014 dan podes 2014.

Variable Penelitian

- ✓ Variabel Respon adalah penduga proporsi tiap kecamatan
- ✓ Variabel penyerta yang diduga mempunyai korelasi dengan tingkat pengangguran dipilih dengan menyesuaikan dengan data Potensi Desa 2011 (PODES 2011) antara lain:
 - 1) Persentase rumah tangga pertanian (sebagai pendekatan faktor usaha informal).
 - 2) Persentase desa dengan keberadaan pasar (sebagai pendekatan faktor usaha informal).
 - 3) Industri kecil.
 - 4) Persentase desa dengan keberadaan SLTA/ sederajat (pendidikan).
 - 5) Rasio penduduk laki-laki.
 - 6) Persentase desa yang terdapat keg dana bantuan usaha non pertanian

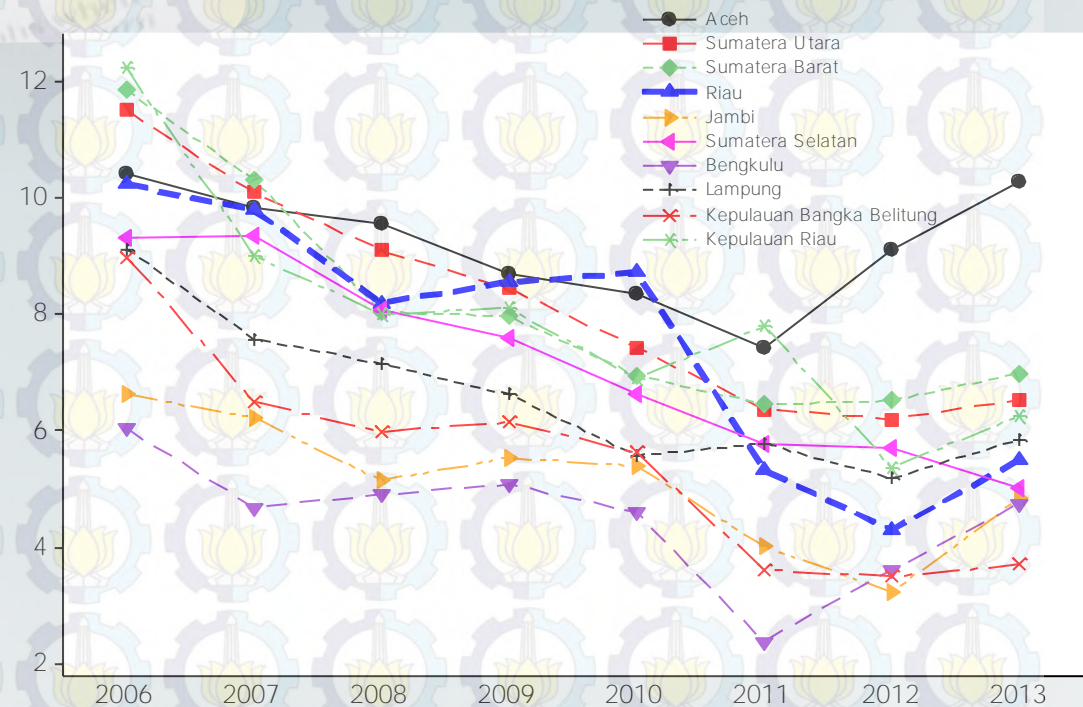
Metodologi (2)

Tahapan Penelitian



Hasil & pembahasan (1)

Gambaran Umum TPT



Gambar Perkembangan TPT Provinsi Provinsi Se-Pulau Sumatera Tahun 2006-2013

Hasil & pembahasan (2)

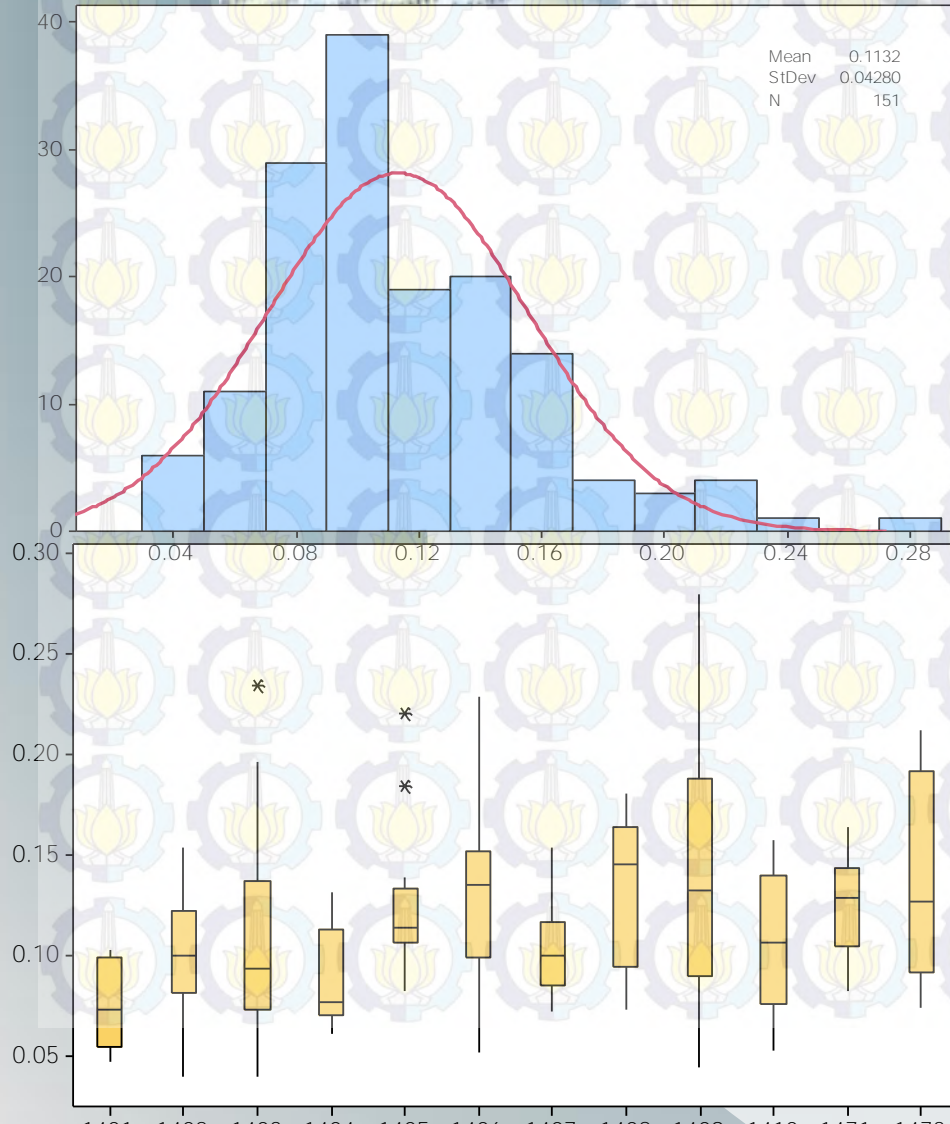
Karakteristik Populasi

Id_Kab	Kabupaten/Riau	Jumlah Kec	Jml RT	Proporsi
1401	Kuantan Sengingi	12	71.002	0,076
1402	Indragiri Hulu	14	86.188	0,104
1403	Indragiri Hilir	20	158.249	0,104
1404	Pelalawan	12	74.772	0,088
1405	Siak	14	89.368	0,118
1406	Kampar	20	160.511	0,130
1407	Rokan Hulu	16	113.974	0,103
1408	Bengkalis	8	113.598	0,143
1409	Rokan Hilir	13	124.667	0,125
1410	Kepulauan Meranti	5	38.271	0,119
1471	Kota Pekanbaru	12	199.569	0,133
1473	Kota Dumai	5	57.217	0,155
1400	Riau	151	1.287.386	0,118

Kode_Kab	Kode_Kec	Proporsi	Kode_Kab	Kode_Kec	Proporsi
1401	1401010	0.094	1402	1402022	0.084
1401	1401011	0.102	1402	1402030	0.094
1401	1401012	0.100
1401	1401020	0.047
1401	1401021	0.066
1401	1401030	0.102	1406	1406020	0.146
1401	1401040	0.071	1406	1406030	0.075
1401	1401050	0.058	1406	1406031	0.229
1401	1401051	0.048
1401	1401052	0.054
1401	1401060	0.087
1401	1401061	0.076
1402	1402011	0.048	1473	1473012	0.171
1402	1402020	0.090	1473	1473020	0.127
1402	1402021	0.126	1473	1473020	0.212

Hasil & pembahasan (3)

Karakteristik Populasi



Normal [#41]					
Kolmogorov-Smirnov					
Sample Size	151				
Statistic	0.10344				
P-Value	0.07352				
Rank	39				
	0.2	0.1	0.05	0.02	0.01
Critical Value	0.08732	0.09953	0.11051	0.12353	0.13257
Reject?	Yes	Yes	No	No	No
Anderson-Darling					
Sample Size	151				
Statistic	1.6821				
Rank	37				
	0.2	0.1	0.05	0.02	0.01
Critical Value	1.3749	1.9286	2.5018	3.2892	3.9074
Reject?	Yes	No	No	No	No
Chi-Squared					
Deg. of freedom	7				
Statistic	12.598				
P-Value	0.08252				
Rank	35				
	0.2	0.1	0.05	0.02	0.01
Critical Value	9.8032	12.017	14.067	16.622	18.475
Reject?	Yes	Yes	No	No	No

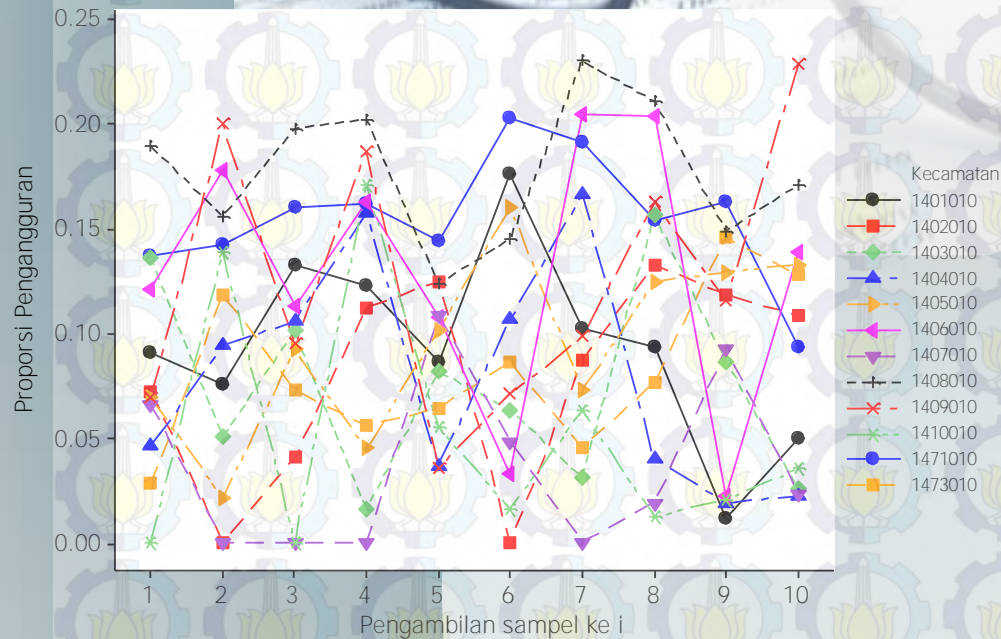
Hasil & pembahasan (4)

Karakteristik Variabel penyerta

Variabel	Rata-rata	Standar Deviasi	Koefisien Variasi	Minimum	Maximum
x1	57,74	25,92	44,89	0,00	94,30
x2	47,24	24,78	52,46	0,00	100,00
x3	2,83	3,88	136,78	0,00	27,27
x4	0,14	0,06	43,92	0,03	0,41
x5	1,05	0,08	7,23	0,88	1,67
x6	39,30	33,02	84,03	0,00	100,00

Variabel	P		logit (P)	
	Korelasi	p-value	Korelasi	p-value
x1	-0,179	0,028	-0,246	0,002
x2	-0,108	0,186	-0,204	0,206
x3	-0,011	0,895	-0,006	0,945
x4	-0,041	0,62	-0,038	0,647
x5	-0,076	0,357	-0,065	0,431
x6	-0,21	0,010	-0,179	0,028

Hasil & pembahasan (5)



Pengambilan sampel

Id_Kab	Kabupaten/Riau	Frame SP2010		Sampel	
		BS	Rata-rata Jumlah RT	BS	RT
1401	Kuantan Sengingi	950	75	35	10
1402	Indragiri Hulu	1096	79	42	10
1403	Indragiri Hilir	1918	82	54	10
1404	Pelalawan	980	76	41	10
1405	Siak	1146	78	42	10
1406	Kampar	2224	72	54	10
1407	Rokan Hulu	1490	76	46	10
1408	Bengkalis	1247	91	46	10
1409	Rokan Hilir	1456	86	46	10
1410	Kepulauan Meranti	544	70	38	10
1471	Kota Pekanbaru	2531	79	54	10
1473	Kota Dumai	614	93	38	10

Id_Kab	Kabupaten/Riau	Rata-rata Proporsi Hasil Sampel	Rata-rata Bias Absolut	Proporsi Populasi
1401	Kuantan Sengingi	0,072	0.013	0,076
1402	Indragiri Hulu	0,100	0.014	0,104
1403	Indragiri Hilir	0,103	0.019	0,104
1404	Pelalawan	0,091	0.012	0,088
1405	Siak	0,117	0.012	0,118
1406	Kampar	0,132	0.017	0,130
1407	Rokan Hulu	0,101	0.024	0,103
1408	Bengkalis	0,149	0.015	0,143
1409	Rokan Hilir	0,123	0.014	0,125
1410	Kepulauan Meranti	0,125	0.007	0,119
1471	Kota Pekanbaru	0,121	0.018	0,133
1473	Kota Dumai	0,161	0.017	0,155

Hasil & pembahasan (6)

Model HB FH

Variabel	Mean (β)	Sd (β)	MC error	2,50%	median	97,50%	Rasio MC Error
X1	-0,00031	0,00033	0,00001	-0,00094	-0,00031	0,00033	0,01663
X2	0,00044	0,00035	0,00001	-0,00023	0,00044	0,00114	0,01631
X3	-0,00008	0,00031	0,00001	-0,00069	-0,00008	0,00055	0,01902
X4	-0,00047	0,00028	0,00000	-0,00102	-0,00048	0,00006	0,01604
X5	0,00031	0,00029	0,00001	-0,00027	0,00031	0,00089	0,01796
X6	-0,00030	0,00029	0,00000	-0,00086	-0,00030	0,00026	0,01582
Konstanta	0,11630	0,00913	0,00016	0,09843	0,11630	0,13420	0,01782

Kode_Kec	Mean (P)	Sd (P)	MC error	2,50%	median	97,50%	Rasio MC Error
1401010	0,09031	0,02937	0,00044	0,03403	0,09047	0,14900	0,01501
1401011	0,07915	0,03495	0,00064	0,00883	0,07968	0,14780	0,01835
1401012	0,09426	0,03488	0,00044	0,02531	0,09431	0,16520	0,01275
1401020	0,06912	0,03353	0,00048	0,00395	0,06913	0,13550	0,01434
1401021	0,06591	0,03219	0,00047	0,00073	0,06645	0,12710	0,01445
1401030	0,04568	0,02924	0,00050	-0,01578	0,04721	0,09894	0,01697
1401040	0,10070	0,02989	0,00049	0,04291	0,10050	0,16070	0,01652
1401050	0,09398	0,02988	0,00056	0,03348	0,09487	0,15010	0,01885
1401051	0,07859	0,03366	0,00051	0,01264	0,07781	0,14710	0,01514
1401052	0,07054	0,03536	0,00052	-0,00113	0,07068	0,13750	0,01466

CI masih mengandung nilai negatif --- kurang tepat

Hasil & pembahasan (7)

Model HB NL & NLrs, BLrs

Variabel	Mean (β)	Sd (β)	MC error	2,50%	median	97,50%	Rasio MC Error
X1	-0,0035	0,0041	0,0001	-0,0119	-0,0034	0,0042	0,0168
X2	0,0055	0,0043	0,0001	-0,0027	0,0054	0,0141	0,0178
X3	-0,0015	0,0038	0,0001	-0,0089	-0,0015	0,0062	0,0180
X4	-0,0063	0,0037	0,0001	-0,0136	-0,0063	0,0008	0,0174
X5	0,0043	0,0037	0,0001	-0,0029	0,0043	0,0119	0,0160
X6	-0,0042	0,0037	0,0001	-0,0121	-0,0042	0,0028	0,0195
Konstanta	-2,1180	0,1095	0,0018	-2,3520	-2,1150	-1,9150	0,0166

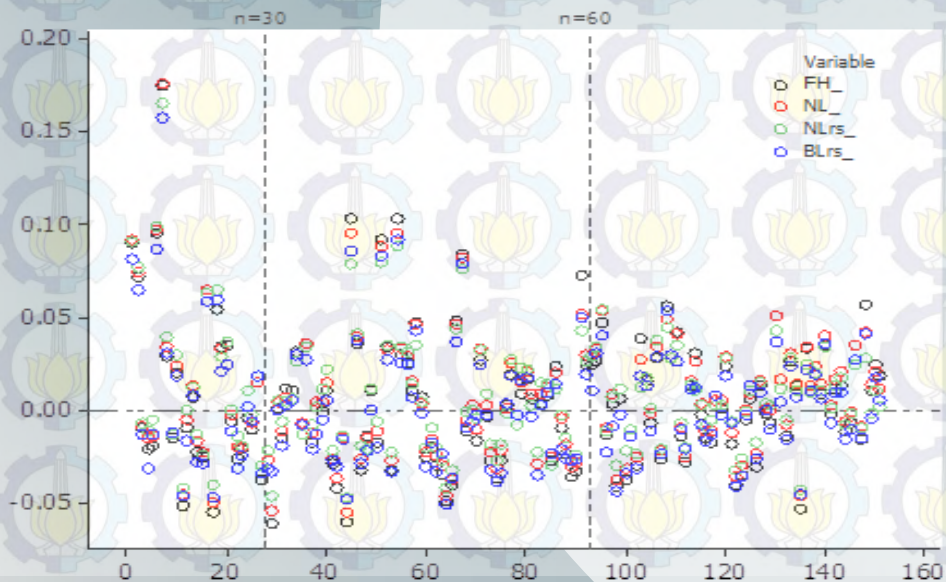
Variabel	Mean (β)	Sd (β)	MC error	2,50%	median	97,50%	Rasio MC Error
X1	-0,0053	0,0041	0,0001	-0,0133	-0,0052	0,0027	0,0133
X2	0,0070	0,0042	0,0001	-0,0013	0,0070	0,01 51	0,0150
X3	0,0002	0,0039	0,0001	-0,0073	0,0002	0,0078	0,0145
X4	-0,0054	0,0035	0,0000	-0,0122	-0,0055	0,0014	0,0138
X5	0,0030	0,0037	0,0001	-0,0043	0,0030	0,0102	0,0156
X6	-0,0057	0,0036	0,0001	-0,0129	-0,0057	0,0013	0,0149
Konstanta	-2,1470	0,1132	0,0016	-2,3800	-2,1430	-1,9370	0,0143

Variabel	Mean (β)	Sd (β)	MC error	2,50%	median	97,50%	Rasio MC Error
X1	-0,0058	0,0038	0,0001	-0,0133	-0,0057	0,0016	0,0139
X2	0,0063	0,0040	0,0001	-0,0016	0,0063	0,0140	0,0170
X3	-0,0005	0,0036	0,0001	-0,0076	-0,0006	0,0064	0,0159
X4	-0,0049	0,0033	0,0000	-0,0112	-0,0049	0,0017	0,0143
X5	0,0036	0,0034	0,0000	-0,0032	0,0036	0,0103	0,0146
X6	-0,0034	0,0033	0,0000	-0,0100	-0,0034	0,0030	0,0139
Konstanta	-2,0990	0,0986	0,0014	-2,2990	-2,0970	-1,9080	0,0139

Hasil & pembahasan (8)

Perbandingan Model

Model	DIC	OAB	OAAD	OAARD	95% fc
FH	-312.78	-0.004	0.0314	0.3087	0.06181
NL	-321.71	-0.007	0.0318	0.3011	0.04194
NLrs	-341.01	-0.007	0.0325	0.3019	0.03532
BLrs	-426.13	-0.002	0.0319	0.3153	0.04783
DE	-	-0.004	0.0601	0.5716	-



Hasil & pembahasan (9)

Prediksi Proporsi Pengangguran Th 2014

Variabel	Mean (β)	Sd (β)	MC error	2,50%	median	97,50%	Rasio MC Error
X1	-0,00607	0,00937	0,00004	-0,02530	-0,00582	0,01166	0,00442
X2	0,00119	0,00865	0,00003	-0,01662	0,00141	0,01769	0,00395
X3	0,00836	0,00853	0,00004	-0,00792	0,00816	0,02583	0,00436
X4	0,01172	0,00842	0,00003	-0,00342	0,01115	0,03005	0,00415
X5	-0,00248	0,00769	0,00003	-0,01872	-0,00211	0,01166	0,00429
X6	-0,00715	0,00758	0,00003	-0,02181	-0,00730	0,00828	0,00452
Konstanta	-3,52900	0,27140	0,00129	-4,14100	-3,50500	-3,06500	0,00477
Sigma	0,95560	0,28270	0,00144	0,52280	0,91560	1,63400	0,00508

$$\text{logit}(P_i) = -3,529 - 0,0061x_1 + 0,001x_2 + 0,008x_3 \\ + 0,012x_4 - 0,0202x_5 - 0,007x_6 + v_i$$

Hasil & pembahasan (10)

Prediksi Proporsi Pengangguran Th 2014

Kecamatan	Proporsi pengangguran	SD	MC error	2.50%	Median	97.50%	Rasio MC error
1401010	0.0292	0.0270	0.0001	0.0016	0.0216	0.1018	0.0044
1401011	0.0548	0.0687	0.0003	0.0027	0.0339	0.2381	0.0043
1401012	0.0968	0.1059	0.0004	0.0055	0.0629	0.4008	0.0040
1401013	0.0840	0.0928	0.0004	0.0050	0.0549	0.3455	0.0043
1401020	0.0798	0.0414	0.0002	0.0285	0.0701	0.1873	0.0042
1401021	0.0918	0.0540	0.0002	0.0288	0.0781	0.2337	0.0039
1401030	0.0614	0.0717	0.0003	0.0039	0.0396	0.2557	0.0043
1401031	0.0780	0.0981	0.0004	0.0031	0.0456	0.3620	0.0042
1401040	0.0301	0.0211	0.0001	0.0051	0.0249	0.0834	0.0044
1401050	0.1001	0.0511	0.0002	0.0362	0.0883	0.2322	0.0046
1401051	0.0677	0.0347	0.0002	0.0236	0.0600	0.1554	0.0044
1401052	0.0438	0.0255	0.0001	0.0114	0.0381	0.1081	0.0045
1401053	0.0547	0.0699	0.0003	0.0022	0.0327	0.2468	0.0040
1401060	0.0480	0.0243	0.0001	0.0159	0.0429	0.1082	0.0044
1401061	0.0478	0.0244	0.0001	0.0154	0.0427	0.1086	0.0039
1402010	0.0842	0.0974	0.0004	0.0042	0.0532	0.3603	0.0046
1402011	0.0650	0.0761	0.0003	0.0039	0.0413	0.2761	0.0040
1402020	0.0400	0.0390	0.0002	0.0022	0.0287	0.1458	0.0046
1402021	0.0360	0.0352	0.0001	0.0045	0.0258	0.1294	0.0041
1402022	0.1135	0.0732	0.0003	0.0319	0.0938	0.3106	0.0042
1402030	0.0197	0.0201	0.0001	0.0008	0.0137	0.0734	0.0046
1402031	0.0754	0.0826	0.0004	0.0051	0.0502	0.3060	0.0044
1402040	0.0943	0.1006	0.0004	0.0068	0.0632	0.3811	0.0044
1402041	0.0691	0.0649	0.0003	0.0078	0.0489	0.2509	0.0042
1402042	0.0442	0.0280	0.0001	0.0125	0.0371	0.1181	0.0043
1402043	0.1096	0.1150	0.0005	0.0074	0.0731	0.4446	0.0046
1402050	0.0164	0.0171	0.0001	0.0007	0.0113	0.0620	0.0044
1402060	0.0789	0.0920	0.0004	0.0039	0.0498	0.3378	0.0043

Kesimpulan & Saran

Kesimpulan

- Penarikan sampel dengan jumlah sampel yang kecil cenderung membuat hasil estimasi langsung proporsi pengangguran kecamatan sangat berfluktuasi dan tidak stabil. Hal ini berbeda jika proporsi pengangguran dihitung untuk tingkat kabupaten. Dengan jumlah sampel yang memadai estimasi langsung untuk tingkat kabupaten cenderung lebih stabil dan secara rata-rata mendekati nilai *true value*-nya.
- Variabel-variabel dari Podes 2011 belum bisa memberikan pengaruh yang signifikan terhadap prediksi proporsi pengangguran tingkat kecamatan. Dengan demikian prediksi pengangguran di kecamatan hanya mendasari pada distribusi proporsi dalam satu provinsi (*random effect*) belum bisa mengakomodir konsep *borrowing strenght* antar kecamatan yang berasal dari variabel penyerta. Hal yang sama juga terjadi pada saat memprediksi pengangguran dengan menggunakan data Sakernas 2014.
- Meskipun model NLrs memiliki nilai DIC yang lebih besar dibanding model BLrs, tetapi pada kasus ini model NLrs terbukti lebih baik untuk memprediksi proporsi pengangguran tingkat kecamatan di Provinsi Riau dilihat dari validasi dengan *Overall Absolute Average Relative Deviation* (OAARD) dan 95% *credible interval* yang tidak mampu mencakup nilai *true value*-nya yang paling kecil diantara model yang lain (95% *credible interval fails to cover true value*).

Kesimpulan & Saran

Saran

- Untuk mendapatkan pemodelan yang lebih baik dalam memprediksi proporsi pengangguran perlu untuk mencari variabel penyerta selain variabel dari Podes. Hal ini untuk mengantisipasi adanya korelasi yang sangat kecil antara variabel podes dengan proporsi pengangguran sehingga dalam model tidak signifikan mempengaruhi prediksi.
- Dalam menyimpulkan model NLrs sebagai model terbaik dibandingkan model FH, NL, dan BLrs masih perlu diteliti kembali mengingat simulasi yang dilakukan hanya sebanyak 10 kali. Dalam penelitian selanjutnya disarankan untuk memperbanyak jumlah simulasi pengambilan sampling serta bisa dicobakan dengan metoda sampling yang lain.
- Perlu untuk membandingkan keempat model HB SAE yang telah dibahas dengan model SAE yang lain seperti model HB dengan tranformasi *arcsine* untuk parameter varian sampling ataupun pendekatan binomial dengan koreksi sampling yang digunakan.
- Untuk melakukan prediksi sebaiknya dilakukan penambahan sampel. Karena prediksi akan semakin baik meskipun variabel penyerta tidak signifikan. Selain itu dengan jumlah sampel yang lebih banyak akan memperkecil kemungkinan suatu daerah tidak terkena (terwakili) oleh sampel.

Daftar Pustaka (1)

- Arora, V. and Lahiri P. (1997), "On the superiority of the Bayesian method over BLUP in small area estimations problems", *Statistical Sinica*, vol. 7, pp. 1053-1063.
- Ubaidillah, Azka. (2014), *Small Area Estimation Dengan Pendekatan Hierarchical Bayesian Neural Network untuk Pemetaan Kemiskinan di Kota Jambi*, Tesis, Institut Teknologi Sepuluh Nopember, Surabaya.
- Bukhori, A.S. (2015), *Pendugaan Area Kecil Komponen Indeks Pendidikan dalam IPM di Kabupaten Indramayu dengan Metode Hierarchical Bayes Berbasil Spasial*, Tesis, Universitas Padjajaran, Bandung.
- Cochran, W.G. (1977), *Sampling Technique*, 3rd ed, Jhon Wiley & Sons, inc., New Jersey.
- Congdon, P.D. (2010), *Applied Bayesian Hierarchical Methods*, CRC Press, Boca Raton.
- Cowles, M.K., dan Carlin, B.P. (1996), "Markov Chain Monte Charlo convergence diagnostics: a comparative review", *Journal of the American Statistical Association*, vol. 91, pp. 883-904.
- Farrell, P.J. (1995), "Hierarchical Bayes approach for estimating small area proportions", *Proceedings of the Survey Research Methods Section*, pp. 63-68.
- Frerich, R.R: Equal Probability of Selection, in Rapid Surveys (unpublised).
- Gelman, Andrew (2006), "Prior Distribution for Variance Parameters in Hierarchical Models". *Bayesian Analysis 1*, vol. 3, pp. 515-534.
- Ghosh, M., dan Rao, J.N.K. (1994), "Small area estimation: an appraisal". *Statistical Science*, Institute of Mathematical Statistics, vol. 9, pp. 55-76.
- Ha, N.S. (2013), *Hierarchical Bayes Estimation of Small Area Means Using Complex Survey Data*. Disertasi P.h.d, University of Maryland.
- Ha, N.S., Lahiri, P., and Parsons, V. (2014), "Methods and Results for Small Area Estimation using Smoking Data from The 2008 Nasional Health Interview Survey", *Staitistics in Medicine*, vol. 33, pp. 3932-3945.

Daftar Pustaka (2)

- Hajarisman, Nusar. (2013), *Pemodelan Area Kecil untuk Menduga Angka Kematian Bayi Melalui Pendekatan Model Regresi Poisson Bayes Berhierarchy Dua Level*, Disertasi, FMIPA IPB, Bogor.
- Hawala, S., and Lahiri, P. (2012), "A Hierarchical Bayes Estimation of Poverty Rates", *Proceedings of American Statistical Association, Section on Survey Research Methods*, Alexandria, VA: American Statistical Association, Forthcoming.
- Jiang, J., Lahiri, P. (2006). "Mixed Model Prediction and Small Area Estimation", *Test*, vol. 15, 1, pp. 1-96.
- Kish, L. (1965). *Survey sampling*, John Wiley, New York.
- Liu, B. (2009). *Hierarchical Bayes Estimation and Empirical Best Prediction of Small Area Proportions*, Disertasi Ph.D, University of Maryland.
- Liu, B., Lahiri, P., and Kalton, G. (2007), "Hierarchical Bayes Modeling of Survey-Weighted Small Area Proportions", *Proceedings of the American Statistical Association, Section on Survey Research Methods*, Alexandria, VA: American Statistical Association, Forthcoming, pp. 3181-3186.
- Lunn, D., Jackson, C., Best, N., Thomas, A., Spiegelhalter, D. (2013), *The Bugs Book; A Practical Introduction to Bayesian Analysis*, CRC Press, Boca Raton.
- Malec, D., Sedransk, J., Moriarity, C., and LeClere, F. (1997), "Small Area Inference for Binary Variables in National Health Interview Survey", *Journal of the American Statistical Association*, vol. 92, pp. 815-826.
- Mohadjer, L., Rao, J.N.K., Liu, B., Krenzke, T., Van de Kerckhove, W. (2012), "Hierarchical Bayes Small Area Estimates of Adult Literacy using Unmatched Sampling and Linking Models", *Journal of the Indian Society of Agricultural Statistics*, vol. 66, pp. 55-63.
- Peters, J., and Eachus, J.I. (1995), "Achieving Equal Probability of Selection under Various Random Sampling Strategies", *Paediatric and Perinatal Epidemiology*, Vol. 9, pp. 219-224.

Daftar Pustaka (3)

- Rao, J.N.K. (2003), *Small Area Estimation*, John Wiley & Son, New York.
- Rumiati, A.T. (2012), *Model Bayes untuk Pendugaan Area Kecil dengan Penarikan Contoh Berpeluang Tidak Sama pada Kasus Respon Binomial dan Multinomial*, Disertasi, Institut Pertanian Bogor, Bogor.
- You, Y. (2008), "An Integrated Modeling Approach to Unemployment Rate Estimation for Subprovincial Areas of Canada, *Survey Methodologi*, vol. 34, pp. 19-27.
- You, Y., and Rao, J.N.K. (2002), "Small Area Estimation Using Unmatched Sampling and Linking Models, *Canadian Journal of Statistics*, vol. 30, pp. 3-15.
- You, Y., and Chapman, B. (2006). "Small Area Estimation Using Area Level Models and Estimated Sampling Variances", *Survey Methodology*, vol. 32, pp. 97-103.
- Xia, H., Carlin, B.P., Waller, L.A. (1995), " Hierarchical Models for Mapping Ohio Lung Cancer Rate", Research Report 95-011, Division of Biostatistics, University of Minnesota.
- Xie, D., Raghunathan, T.E., Lpkowski, J.M. (2006), " Estimation of The Proportion of Overweight Individuals in Small Areas-a robust extension of The Fay-Herriot Model", *Statistics In Medicine*, vol. 26, pp. 2699-2715.

Terimakasih



abuiffah



Arip_jy@Twitter



#arip_jy



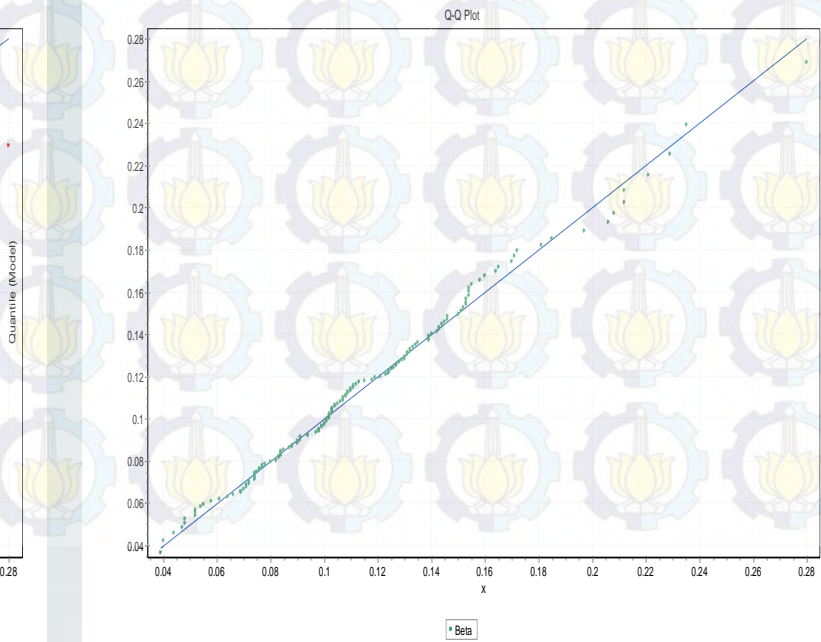
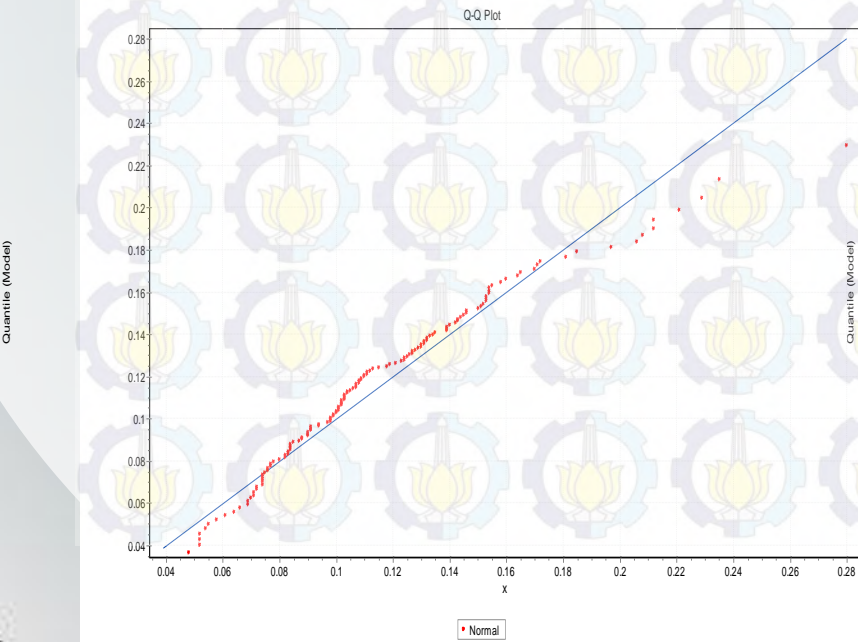
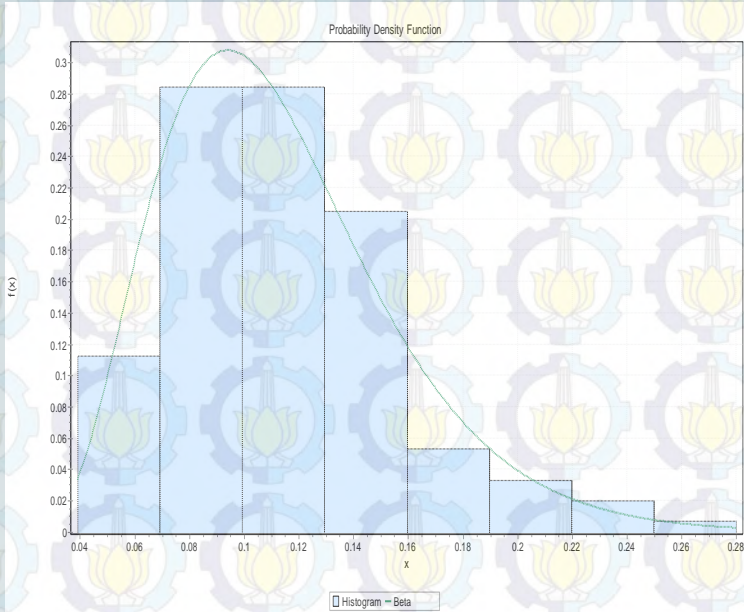
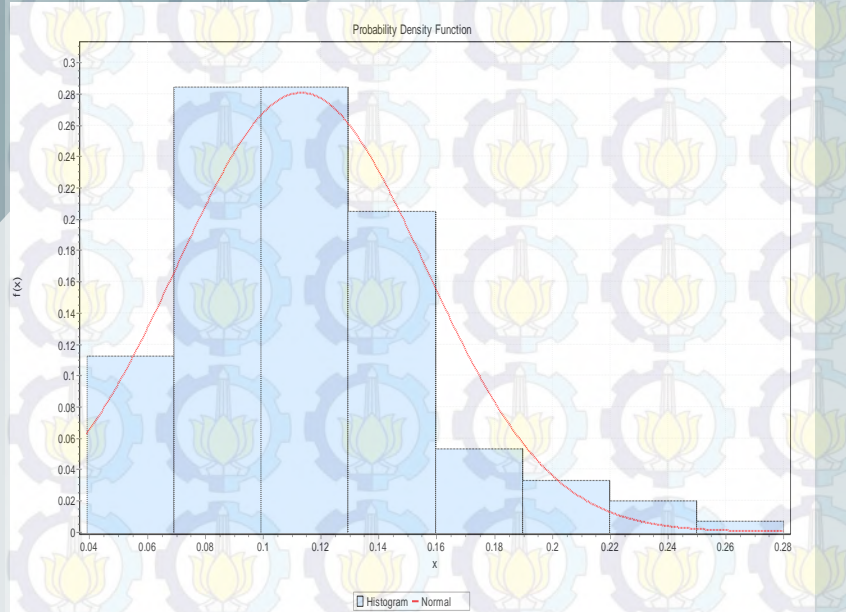
itsar82@gmail.com

Uji distribusi proporsi (Normal & Beta distribusi)

Normal [#40]					
Kolmogorov-Smirnov					
Sample Size	151				
Statistic	0.10344				
P-Value	0.07352				
Rank	37				
α	0.2	0.1	0.05	0.02	0.01
Critical Value	0.08732	0.09953	0.11051	0.12353	0.13257
Reject?	Yes	Yes	No	No	No
Anderson-Darling					
Sample Size	151				
Statistic	1.6821				
Rank	36				
α	0.2	0.1	0.05	0.02	0.01
Critical Value	1.3749	1.9286	2.5018	3.2892	3.9074
Reject?	Yes	No	No	No	No
Chi-Squared					
Deg. of freedom	7				
Statistic	12.598				
P-Value	0.08252				
Rank	34				
α	0.2	0.1	0.05	0.02	0.01
Critical Value	9.8032	12.017	14.067	16.622	18.475
Reject?	Yes	Yes	No	No	No

Beta [#1]					
Kolmogorov-Smirnov					
Sample Size	151				
Statistic	0.04631				
P-Value	0.88729				
Rank	12				
α	0.2	0.1	0.05	0.02	0.01
Critical Value	0.08732	0.09953	0.11051	0.12353	0.13257
Reject?	No	No	No	No	No
Anderson-Darling					
Sample Size	151				
Statistic	0.3264				
Rank	17				
α	0.2	0.1	0.05	0.02	0.01
Critical Value	1.3749	1.9286	2.5018	3.2892	3.9074
Reject?	No	No	No	No	No
Chi-Squared					
Deg. of freedom	7				
Statistic	7.4156				
P-Value	0.38693				
Rank	20				
α	0.2	0.1	0.05	0.02	0.01
Critical Value	9.8032	12.017	14.067	16.622	18.475
Reject?	No	No	No	No	No

Uji distribusi proporsi (Normal & Beta distribusi)



Sampel	Model FH					Model NL			
	Dbar	Dhat	pD	DIC		Dbar	Dhat	pD	DIC
1	-365.31	-401.82	36.50	-328.81		-374.44	-411.54	37.11	-337.33
2	-363.35	-397.91	34.55	-328.80		-373.38	-410.02	36.64	-336.75
3	-370.47	-402.83	32.37	-338.10		-376.05	-408.54	32.49	-343.56
4	-336.45	-372.84	36.39	-300.06		-347.57	-384.41	36.85	-310.72
5	-324.89	-363.92	39.03	-285.86		-334.93	-376.05	41.12	-293.81
6	-335.72	-362.91	27.19	-308.53		-342.93	-371.64	28.71	-314.22
7	-335.22	-367.99	32.77	-302.45		-352.18	-390.23	38.05	-314.14
8	-339.91	-381.68	41.77	-298.13		-355.38	-397.73	42.35	-313.02
9	-328.64	-363.80	35.16	-293.48		-338.94	-375.34	36.40	-302.54
10	-380.79	-418.02	37.23	-343.56		-388.98	-426.96	37.98	-351.00
Rata-rata	-348.07	-383.37	35.30	-312.78		-358.48	-395.25	36.77	-321.71

Sampel	Model NLrs					Model BLrs			
	Dbar	Dhat	pD	DIC		Dbar	Dhat	pD	DIC
1	-405.94	-457.42	51.48	-354.46		-475.96	-528.90	52.94	-423.02
2	-404.87	-457.71	52.84	-352.03		-494.72	-532.33	37.61	-457.11
3	-400.17	-443.51	43.34	-356.83		-486.20	-541.37	55.17	-431.02
4	-378.33	-429.77	51.45	-326.88		-440.73	-492.72	51.99	-388.75
5	-377.73	-436.58	58.86	-318.87		-477.02	-537.51	60.50	-416.52
6	-378.15	-425.42	47.27	-330.88		-482.67	-524.85	42.18	-440.49
7	-393.94	-449.96	56.02	-337.92		-461.80	-509.65	47.86	-413.94
8	-393.22	-450.96	57.74	-335.48		-473.09	-523.65	50.56	-422.54
9	-384.74	-439.51	54.77	-329.97		-482.35	-530.74	48.39	-433.96
10	-416.37	-465.91	49.54	-366.84		-482.35	-530.74	48.39	-433.96
Rata-rata	-393.35	-445.68	52.33	-341.01		-475.69	-525.25	49.56	-426.13